The perceptual origins of the abstract same/different concept in human infants

Journal Article

Citation:

Addyman, C.; Mareschal, D. (2011)
The perceptual origins of the abstract same/different concept in human infants – *Animal Cognition* 13(6), pp.817-833

Animal Cognition manuscript No.
(will be inserted by the editor)

The perceptual origins of the abstract Same/Different concept in human infants.

Caspar Addyman · Denis Mareschal

Received: date / Accepted: date

Abstract Very few experiments have studied the two item Same/Different relation in young human infants. This contrasts with an extensive animal literature. We tested young infants with two novel tasks designed specifically to provide convergent comparative measures. Each infant completed both tasks allowing an assessment of their understanding of the abstract concept rather than task-specific abilities. In a looking time task with photographic stimuli we found that 8-month-olds are sensitive to the relation but 4-month-olds are not. The second task used an anticipatory eye movement paradigm with simple geometric stimuli. On each trial, two colored shapes appear and moved upwards behind an occluder. They reappeared on either the upper left or right depending on the relation between them. Infants at both ages learned and generalized the dependency but only for the different relation. These results

Support for this research comes from European Commission Framework 6 NEST contracts 029088 and 516542.

Centre for Brain and Cognitive Development, Birkbeck College, University of London, Malet Street, London, WC1E 7HX, UK
Tel.: +44-207-0790756, Fax: +44-207-6316587 E-mail: c.addyman@bbk.ac.uk

show that human infants can learn the Same/Different concept but that, in strong continuity with animal results, their abilities are firmly grounded in perception.

(164 words)

Keywords Same/Different · Human Infants · Concept Learning

1 Introduction

A fundamental characteristic of human cognition is the ability to make use of abstract concepts and perceive similar relations between otherwise unrelated items. These can take a very wide variety of forms including judgements of numerosity and of comparative quantity and size (*e.g., largest, smallest*), relations expressed by spatial prepositions (*e.g., above, behind, between*) and numerous others. One of the simplest examples is the two item same/different (S/D) distinction. Consider the set of 9 distinct items A to I and the pairs AA, BB, CC, DE, FG, HI. The first three pairs share the relationship of sameness and (by contrast) the last three have the common property of difference. These relationships hold at an abstract level, irrespective of the actual items used and it is this abstraction that defines the relationship. Despite the apparent simplicity from an adult perspective, it has been difficult to demonstrate conclusively with species other than higher mammals and largely ignored in the human infancy literature. Developmental and comparative approaches both have a lot to reveal about adult human concepts (Mareschal et al., 2010).

Therefore, we sought to investigate S/D concept learning in infants, taking into account the extensive animal learning literature.

Until recently the received wisdom was that humans find S/D tasks very easy and animals find them very difficult. For humans the S/D distinction would be expected to be a sharp, definite qualitative difference whereas for animals, if they succeed at all, the difference would always be graded, uncertain and quantitative. The difference between animals and humans is portrayed as the contrast between a perceptual, stimulus-driven process and a genuinely abstract concept ability. For example, Penn et al. (2008) present the S/D concept as an example of a qualitative difference between animal and adult human cognitive abilities. In contrast, Wasserman and colleagues suggests that human and animal abilities exist on a continuum that takes account of the number of items to be judged as same or different and the similarity between them (see Wasserman and Young, 2010 for a recent review).

Evidence from studies with the human adults is equivocal. Smith et al. (2008) find a sharp divide between humans and rhesus monkeys on a task that systematically manipulated the similarity of pairs of random polygons. Using signal detection theory, they found that humans assigned a strict, rule-like criteria for Sameness whilst the monkeys performance was more graded. Making the instructions slightly more vague to reduce the influence of language made human performance slightly less certain but still apparently conceptual. Whereas, a series of studies by Wasserman and colleagues (Young and Wasserman, 2001, 2002b; Castro et al., 2006) show that adult humans some-

times respond to multiple item redundancy and between-item similarity as if it were a continuously changing value, similarly to pigeons' responce to entropy (Young and Wasserman, 2002a and see below). However, the majority of human participants those studies responded in a categorical fashion suggesting there may be a special human concept of S/D over and above the perceptual regularities that animals appear to exploit. Therefore the gap between adults and animals remains large. . It is therefore instructive to consider the development these abilities in human infants. This may bridge the gap between animals and humans. By testing infants with paradigms comparable to the animal literature, we may reveal the sudden appearance or innate presence of a uniquely human concept or we may find evidence that human perceptual skills sharpen and mature. Additionally, the animal literature makes it clear that it is important how the question is asked. The abilities of animals are highly task dependent and different tasks can be reveal different distinctions between species. These contrasts are often overlooked when abstract concepts are ascribed to young infants and suggest that multiple paradigms should be used to establish a genuinely abstract ability. We begin by surveying the animal literature before considering any existing work with human infants that bears on this question.

1.1 Same/Different learning in animal subjects

For a long time, animal researchers had difficulty demonstrating S/D discrimination in pigeons or rats. Zentall and Hogan and colleagues (Zentall and

Hogan, 1974, 1976; Zentall et al., 1981) demonstrated that pigeons could use aspects of stimulus similarity and difference in a match-to-sample task. For example, in Zentall and Hogan (1974) birds were trained on the color match or mismatch between cue and target stimuli (e.g., a red or green cuing lamp matched/mismatched to red or green target lamps) and were tested for their ability to transfer this to new stimuli (e.g. blue and yellow lamps). Speeded transfer was found when the second task was congruent to the first. However, Premack (1983) pointed that the match-to-sample technique does not necessarily provide a good test of animals ability to use the abstract relational properties of the stimuli because they may simply be responding to the items they have seen before. Indeed, even honey bees have been shown to succeed on S/D with a delayed match to sample technique (Giurfa et al., 2001).

Premack (1983) proposed that generalization in the two (or more) item S/D task is better suited to testing genuinely abstract S/D concept learning. This has proved difficult with animals. Pearce (1991) found that with very extensive training pigeons could learn to discriminate when relation was of equality or difference in the height of two bars. However, pigeons had failed on an earlier version of the task (Pearce, 1988), leading him to be believe the pigeons were simply memorizing all the possible configurations. This conflict between item-specific learning and acquiring a generalized response is a common problem in pigeon categorization experiments (Wright, 1997) The pigeons' behavior will always be some trade off between the two. Therefore the crucial measure of performance will be the response to completely novel stimuli.

Blaisdell and Cook (2005) reported a successful demonstration of two item S/D learning in pigeons where the learning stimuli were colored geometric shapes and the pigeons were tested on transfer to novel colored shapes. The training stimuli were taken from a set of six shapes in six different colors. On a single training trial a same pair (e.g., two red squares) and a different pair (e.g. a purple chevron and a green star) were displayed on either side of computer monitor and half the pigeons were reinforced to peck the same pair and half to peck different pair. Six out of six pigeons reached an 80% criterion within 28 training sessions. Pigeons were then tested on transfer to either novel shapes, novel colors or both novel. Transfer occurred in all conditions but was weakest in the novel color condition and strongest when both cues were changed. This provides strong evidence that pigeons can transfer in both dimensions, though performance was driven primarily by color and the dimensions seemed to be treated independently. Thus it seems that although pigeons can make some abstraction based on the S/D relation this seems to be tied to individual dimensions rather than taking place at a more 'conceptual' level.

Further evidence in support of a relatively low-level perceptual account of pigeons abilities comes from Young and Wasserman (2002a). In their experiment, pigeons responded to a 16 item array of computer icons that varied in entropy from low (all same) to high (all different), responding was measured in a go/no-go task where arrays with higher or lower entropy were selectively reinforced. Pigeons trained to respond positively to high entropy arrays showed better discrimination. Young and Wasserman provided an analysis

which supported the interpretation that the pigeons multi-item S/D judgements followed a logarithmic (rather than linear) entropy function, suggesting that same-difference could be considered a dimensional rather than a categorical distinction in certain cases. However, that does not prevent there from also being an abstract concept particularly for pairwise S/D. Flemming et al. (2007) demonstrated not only that a larger array (higher entropy) facilitated learning of S/D relations in rhesus monkeys. But also that they could subsequently succeed on a previously failed 2-item relation utilising a colour cue to label represent the 'concept' of S/D. Interestingly, these same monkeys then went on to fail a slightly higher order task assessing 'relations between relations' (Exp. 3, Flemming et al., 2007).

Recently, Wright and Katz (2006) ran a comparative study with pigeons, rhesus monkeys and capuchin monkeys using sets of photographic stimuli which provided strong evidence of learning and generalization in all species. Rhesus and capuchin monkeys learned at comparable rates while pigeons learned more slowly. The animals were required to give one response if the pictures matched and another if they were different. Correct responses resulted in a food reward. All species reached a criterion of 80% with a training set of just 8 items (giving 8 same item pairs and 56 different item pairs). However, no animals initially learned to transfer this to novel test items. Increasing the size of the training in stages lead to full generalization (when accuracy with novel items reached equivalent accuracy as baseline) at 128 items for the monkey species and 256 items for pigeons. Thus there were quantitative but no

qualitative differences between species and a strong criterion for abstractness (equivalent performance with novel material) was met. Wright and Katz speculated (based on human similarity ratings) that this continuity would extend to higher mammals. This appears to be supported by Mercado et al.'s (2000) demonstration of S/D learning in two bottlenose dolphins, albeit using a very different paradigm.

Likewise, in a widely cited study, Oden et al. (1990) tested 11 month old infant chimpanzees with several two item S/D tests. In one experiment, the chimpanzees were given five minute sessions with pairs of toys to physically examine. The two toys could either be the same or different from each other (e.g. AA or CD). Immediately afterwards the chimpanzees were given a novel pair of toys that maintained or changed the relationship. They looked longer to the novel relationship. However, in another experiment the chimpanzees failed on a physical matching to sample task where they were required to match by hand a pair of items to another pair demonstrating a similar relationship. In initial training, the chimpanzees succeeded in reaching a 10/12 criterion but they could not transfer this to novel test sets. This contrasts with the monkeys and pigeons in Wright and Katz (2006). Oden et al. speculate that it was the instrumental aspect of the task that led to the failure with the infant chimpanzees unable to integrate perceptual awareness with motor planning.

Many of these comparative results are of interest from a human developmental perspective. The success of a wide range of species at generalizing the S/D concept suggests that human infants ought to be able both to per-

ceive and act on the abstract relational information and that, at least in the case of perception, there is no reason to suspect this ability will be present early on. Furthermore, infants might be expected to perceive the S/D relation when tested with either rich photographic (Wright and Katz, 2006) or impoverished geometric stimuli (Blaisdell and Cook, 2005). However, Young and Wasserman's (2002) study shows that care must be taken in giving too rich a conceptual interpretation to S/D discrimination. Blaisdell and Cook (2005) are careful to emphasize the importance of generalization as a measure of a genuinely abstract concept. While chimpanzee's failure to perform physical matching (Oden et al., 1990) emphasizes the difference between perceptual awareness and instrumental knowledge. It suggests that lack of integration between developing systems may lead to failures on certain tasks. The overall message from the comparative literature is that one might expect infants to succeed on a range of S/D tasks and demonstrate genuine abstract S/D concept providing care was taken with the design and demands of the task. The next section looks at previous relevant studies with children and infants.

1.2 Same/Different learning in human infants

Only two studies (Tyrrell et al., 1991, 1993) have directly investigated the S/D concept with human infants. An earlier study (Caron and Caron, 1981) found that infants habituated to matching pairs of items dishabituated to non-matching pairs. However that study was not designed specifically to look at S/D concept learning and because infants were only ever habituated to

sameness, non-conceptual explanations could not be ruled out, such as the infants responding to the greater complexity or the broken symmetry of the novel Different display.

Tyrrell et al. (1991) designed their study to address this problem. Using a similar design to Experiment 2 of Oden et al. (1990), Tyrrell et al. tested 29 week old infants using a novelty preference test. Each infant took part in two familiarization/novelty preference tests, one concrete and one abstract. In the concrete test, the infants were familiarized with pairs of small toys affixed to a board, infants in the Same condition had two 20s exposures to a pair of matching toys (AA) and infants in the Different condition saw two dissimilar toys (BC). All infants then saw pairs AA and BC next to each other and in both groups they preferred the unfamiliar pair. In the second abstract phase of the experiment, the infants in the two groups were familiarized to a new pair of items with appropriate relation (DD or EF respectively) and tested for preference for new items GG vs HI. Infants in both groups looked significantly longer at the pair that had a novel relation relative to familiarization. There were no interactions between conditions, although there was a main effect that infants in the Same condition looked longer overall. The fact that infants in both conditions responded to the novel relation in the abstract test suggests that they are responding at a conceptual level. However, the very limited familiarization phase showing just one pair of items in each case and using 'toys' constructed out of a common set of parts increase the likelihood that

the response was to some concrete similarity across conditions than through the use of a completely abstract S/D concept.

Tyrrell et al. (1993) used a conditioned head-turn paradigm with a pair of matching toys on one side and a pair of mismatching toys on the other. Turning to the correct side caused a recording of a children's story to start playing. Infants were assigned to one of four conditions and were reinforced over three separate blocks of four, four and eight trials respectively, with an new set of objects on each block. Infants in one condition were always reinforced for looking at the same pair, whilst in another condition were always reinforced for looking at the different pair. Infants in the two other conditions were reinforced in the same way for blocks 1 and 2 but were then switched to for the final block to be reinforced for the opposite relation. The results supported the hypothesis that the infants were learning the concept rather than simple association to particular items. All groups increased their looking to the correct side over the first 2 blocks but the infants in the switch conditions were impaired at the start of the 3rd block. Tyrrell et al. (1993) argued that this pattern of results mirrors the reversal shifts found in discrimination learning and fits the predictions of the House et al. (1974) model of discrimination learning. They also consider the conditioned head turn paradigm to be an instance of instrumental learning and so drew a contrast between the success of human infants in their task and the failure of the infant chimpanzees in Oden et al. (1990), although they acknowledge that the different task demands render it impossible to make a direct comparison.

An awareness of similarity and difference relations between items is a necessary pre-requisite for success in the artificial grammar tasks of Marcus et al. (1999) In their Experiments 1 and 2 it is also a sufficient condition. Here 7.5-month-olds were habituated to an artificial grammar of syllables of synthesized speech. Infants who had been exposed to several minutes of an ABB tri-syllabic pattern showed a preference for ABA patterns over ABB patterns when tested with completely novel syllables. An awareness of the presence of a reduplicated syllable would be sufficient to discriminate between the two types of grammar. In Marcus et al. (1999) Experiment 3, the contrast was between an AAB and ABB pattern and a simple S/D discrimination would no longer be sufficient. Equally informative are Marcus et al. (2007) where the AAB/ABA/ABB discrimination was not learned directly with non-speech stimuli but could be transferred from speech stimuli, and Johnson et al. (2009), where 8- and 11-month-olds were tested with looming geometric shapes and more easily learned grammars with late item repetition (ABB) than early repetition (AAB) or non-adjacent repetition (ABA). These studies suggest that S/D discrimination in human infants might be a graded phenomenon just as it is in other animals.

Saffran et al. (2007) replicated the Marcus et al. (1999) auditory findings in the visual domain using a task analogue of that task using pictures of dogs and cats arranged in triplets of dog-dog-cat, etc. Saffran et al. contrasted infants success with pictures of dogs and cats to their failures with simple geometric shapes and musical tones (Marcus et al., 2007, Experiment 1) and speculated

that the stimuli must themselves form familiar categories before infants are able to map relations from one case to another. It is an open question as to the extent to which S/D discrimination is a more general process or relies on pre-existing categorical knowledge.

The existing human infant literature has only investigated the abstract S/D concept superficially and indirectly. Yet S/D is the simplest example of a genuinely abstract relation and it is, therefore, of substantial interest to know the limits and origins of infants abilities on this task. For example, no research has looked for asymmetries between Same and Different highlighted by Smith et al. (2008). Moreover, it is notable that no studies report any *failures* of younger infants to respond to S/D relations. Extensive research within category learning (e.g. Eimas and Quinn, 1994) and spatial perception (e.g. Quinn, 2003) has shown that these abilities become increasingly sophisticated with age. It would of interest to know more about the development of infants' abilities with more abstract relations, such as S/D.

The present study has several aims. The primary objective is to test pre-linguistic human infants using convergent experimental designs that probe the robustness of infants' abstract S/D conceptual knowledge to the same stringent level as is required of other species. In particular, it is hoped the infants will meet Blaisdell and Cook's (2005) criteria of equal performance with unfamiliar materials. An additional aim is to demonstrate continuity with the animal literature using tasks that have some common elements (allowing for necessary differences in experimental design required for testing such different

populations.) Thus, the experiments will use both rich photographic stimuli and simplified colored, geometric shapes. The design will include a perceptual and an instrumental task to explore the differences that were observed in Oden et al. (1990) but which were not found in the work of Tyrrell and colleagues. Additionally, the study aims to address the questions raised by Saffran et al. (2007) about whether pre-existing category knowledge and/or domain specificity are important for S/D learning in human infants. Finally, as ought to be the case with any infancy study, we are interested in the development of these abilities. How early do competencies arise and how do they change over time?

Taking all these factors into consideration, two complimentary tasks are presented that look at S/D learning. One task is a simple habituation/dishabituation paradigm using photographic stimuli (see Mareschal and Quinn, 2001). The items are unrelated to each other, not forming an particular category and infants will be familiarized with either a same or a different condition as a between subjects variable. A second task adapts McMurray and Aslin's (2004) anticipatory eye movement paradigm to assess infants' learning of associations between Sameness leading to one outcome (e.g., an occluded pair of objects reappears on the left) and Difference leading to another outcome (e.g., objects reappear on the right). Here, the stimuli will be simple geometric shapes and both learning and generalization will be assessed. The requirement for the infant to actively anticipate the correct side makes this a task that investigates instrumental learning. The same infants are tested on both tasks. This pro-

vides a measure of the robustness of the abstract S/D concept at the level of the individual.

2 Experiment 1 - S/D Discrimination using Photographic Stimuli in a Looking Time Task

Experiment 1 uses a standard habituation/dishabituation paradigm to test for sensitivity to the abstract S/D relation in four and eight month old human infants. The success of 7.5-month-olds in the related tasks of Tyrrell et al. (1991) and Marcus et al. (1999) suggests that by 8 months infants should succeed on a direct test of S/D discrimination. An obvious question is if these abilities are present in younger infants. Caron and Caron (1981) found that infants as young as 12 weeks old dishabituated to *Different* items after being familiarized to items showing the *Same* relationship. However by not running the converse condition, that experiment could not rule out explanations in terms of symmetry or complexity. Therefore, a group of four month olds were also tested.

A set of unrelated photographic stimuli is chosen to give a comparison to the work of Wright and Katz (2006) whilst being able to address Saffran et al.'s (2007) hypothesis that infants will not succeed on such tasks unless the objects come from a familiar or coherent category.

2.1 Method

2.1.1 Participants

A total of 30 full-term human infants provided data, half were around 8 months old (M= 249 days, range 232-259 days, 6 female, 9 male) and half were around 4 months old (M= 124 days, range 117-129 days, 9 female, 6 male). A further seven infants were tested but were excluded due to fussiness (5) or experimental error (2).

Infants were recruited via the centre's participant database. Parents were not paid for their infants' participation but infants were given a small present and travel expenses were reimbursed. All infants took part in both Experiment 1 and Experiment 2 in a single visit. The order of experiments was counterbalanced.

2.1.2 Apparatus

A Macintosh G4 computer running MATLAB R2006a with the PsychToolBox (Brainard, 1997; Pelli, 1997) controlled display of the stimuli and monitored the habituation criterion, using experiment control routines written by the authors. The stimuli were displayed on a Samsung 42" plasma screen and infants reactions were filmed and recorded using an infrared DV camera for later offline coding.

2.1.3 Stimuli

The stimuli were drawn from a set of 25 colored photographs of inanimate objects. All the pictures were scaled to be approximately the same size (300x300 pixels). Pairs of items were displayed on a uniform mid-grey background at a screen resolution of 768 x 1080 pixels with their centers 520 pixels apart, which equated to a visual angle of around 10 ° per item at a separation of around 32 °.

2.1.4 Procedure

Infants were seated on their care-giver's lap approximately 1.3m in front of the monitor with the middle of the monitor at eye-level. The care-giver was instructed not to interact with their infant during the experiment. The experimenter was seated at a control desk in the same room as the infant but hidden behind a heavy curtain. The lighting was kept low and quiet classical music was played in the background throughout the experiment. When the infant was settled, the familiarization phase began.

The screen started as a blank grey slide and each trial began with a 'boing' sound effect (approx. 74dB). Figure 1 gives a schematic representation of the procedure. Pairs of items then appeared on the screen and remained there whilst the infant continued to look at them. A trial finished if infants looked away for a single 500ms period, if they accumulated 1000ms of glances away, or if a maximum of 20s had elapsed. Looking behavior was coded on-line by an experimenter holding a key down on the computer whenever the infant

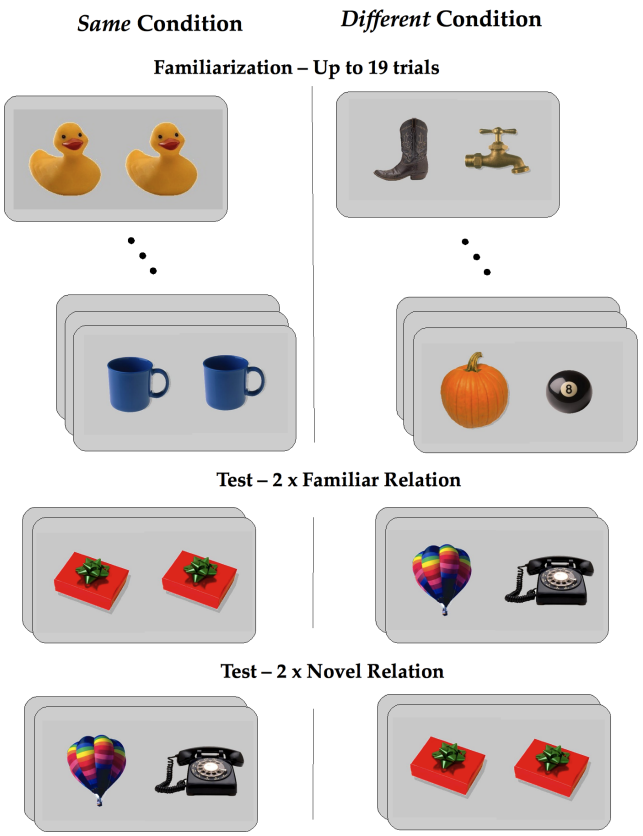


Fig. 1: Procedure for Experiment 1. Infants were assigned to either the *Same* or *Different* condition. They were familiarized to pairs of items demonstrating the relevant relation and then shown four novel test trials. The first two had the familiarized relation while the last two had the opposite relation. A yoked design ensured that pairs of infants in each condition saw the same test images but in opposite orders.

was watching the screen. The control software determined when a look away criterion or a maximum time was reached. The familiarization phase ended if the average length of last two trials was 50% less than the average of the first two trials. Otherwise, there were a maximum of 19 familiarization trials. These were followed by 4 test trials.

The stimuli displayed depended on the experimental condition. Infants were randomly assigned to either the *Same* or the *Different* condition. For infants in both conditions, each trial featured a novel pair of items. The order and selection of items determined by a pseudo-random permutation and care was taken that infants in the different condition did not see more images overall. The experiment employed a 'yoked' design, so that the same sets of items for one member of each group. A random subset of 19 items where used in the familiarization trials. For infants in the *Same* condition each stimulus consisted of an item from this set paired with itself. For infants in the *Different* condition the item on the left was always paired with different item on the right. The item that appeared left position on the n-th trial would re-appear on the right position on trial n+7 (modulo 19). For each infant, the remaining 6 pictures were used to produce the test items, two *Same* test pairs and two *Different* test pairs. An infant in the *Same* condition would see the *Same* test items first then the *Different* test items, a yoked infant in the *Different* condition would see the identical test images but in the opposite order.

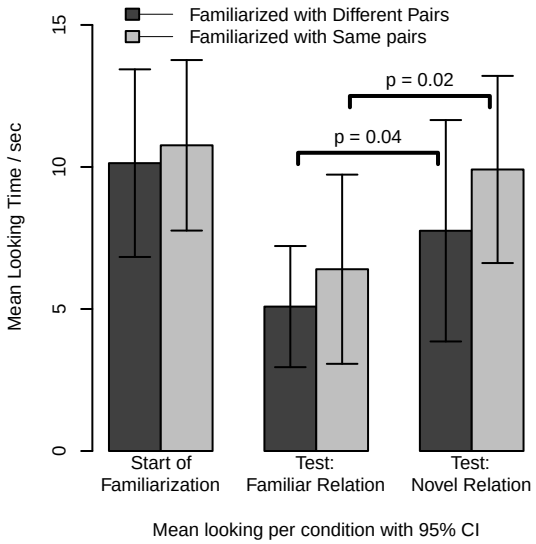
2.2 Results

Looking time scores were calculated by off-line coding of the video recordings. A second experimenter, blind to the experimental hypothesis, double coded the videos of a randomly selected 20% of the infants. A Pearson correlation on the two sets of data gave a high degree of inter-rater reliability, $r = 0.96$, $p = 0.001$. The mean looking times grouped by age are summarized in Figure

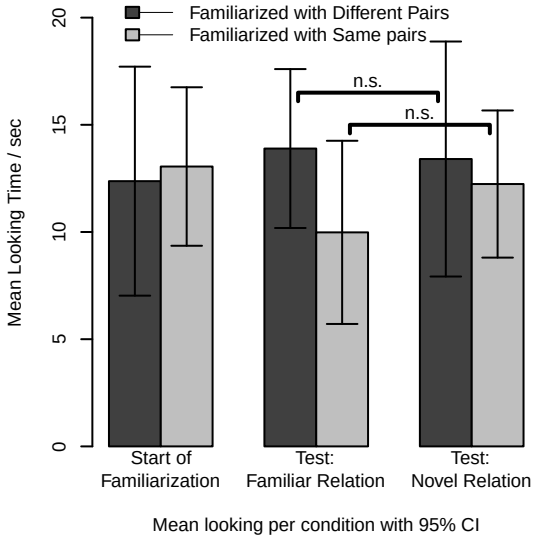
2. For each infant the average of their looking times for first two familiarization trials (Start of familiarization) was compared to the average for the two test trials with the same relation as in familiarization (Test: Familiar Relation) and for the two novel relation trials (Test: Novel Relation), with the comparison between the two test conditions giving an indication of discrimination of the relation.

It was notable that very few infants habituated. Only 4 out of 15 of the 8-month-olds and 5 out of 15 of the 4-month-olds met the habitation criterion of a 50% drop in looking. The 8-month-olds saw an average of 18.5 trials during familiarization and 4-month-olds saw an average of 18.3. For this reason, habituation was not included as a factor for analysis and we looked at all infants together. The experiment was concerned with whether there were age differences in response pattern and whether infants significantly increased their looking to the test trials showing the novel abstract relation.

Therefore, a mixed 3x2x2 ANOVA was conducted with trial type as the within subject variable (*Start of Familiarization, Test: Familiar Relation, Test: Novel Relation*), and condition (*Same* or *Different*) and age *(8 months old* or *4 months old)* as the between subject variables. This revealed a main effect of age, $F_{1,26} = 7.96$, $p = 0.009$, indicating a difference in patterns of responses between 4- and 8-month-olds. As expected, there was a main effect of trial type, $F_{2,52} = 0.046$, which was explored with planned comparisons (see below). Crucially, there was no main effect of condition, $F << 1$, nor



(a) 8-month-olds



(b) 4-month-olds

Fig. 2: Mean looking times in Experiment 1. The light and dark bars indicate infants in the *Same* and *Different* conditions respectively. The leftmost set of bars show average looking at the start of familiarization, while the middle and rightmost bars show responses to familiar and novel test conditions. Planned comparisons and 95% confidence intervals are shown.

were there any interactions all F's < 1, indicating that infant were responding equivalently in both the *Same* and *Different* conditions.

Planned comparisons investigated the infants response to the S/D relation. At 8 months old, infants did show a significant decrease in looking to the familiar relation, $t_{14} = 3.51$, $p = 0.003$, and a subsequent increase in looking to the novel relation, $t_{14} = 3.56$, $p = 0.003$. This was found for both conditions. Eight month old infants who had been familiarized with Same pairs looked longer at the Different test items compared to the Same test items, $t_7 = 2.78$, $p = 0.02$. While 8-month-olds who saw Different pairs initially looked relatively longer at test pairs with two Same items, $t_6 = 2.08$, $p = 0.04$. However, the 4-month-olds showed no overall decrease, $t_{14} = 0.44$, $p = 0.33$, no recovery $t_{14} = 0.65$, $p = 0.26$ nor any decrease or recovery within condition, all t's < 1.

2.3 Discussion

These results provide strong evidence that 8 month old human infants are sensitive to the abstract relation of *S/D*. Infants who had been familiarized to pairs of same or different objects showed a significant increase in looking when shown pairs of new objects with a novel relation. This supports previous findings that 7.5-month-olds succeed on artificial grammar tasks (Marcus et al., 1999; Saffran et al., 2007) and that 7-month-olds can perform some simple S/D discriminations (Tyrrell et al., 1991, 1993). The study also provides continuity with the findings of Wright and Katz (2006) showing that human infants are also able to make a S/D discrimination using photographic stimuli. This

continuity with the animal results and the strength of the effect in the older infants suggests that the mechanism is more general than previously believed.

Saffran et al. (2007) speculate that an awareness of category membership was a pre-requisite for noticing higher level commonalities in infants. But in this case there was no specific perceptual or conceptual category that all objects belonged to. Yet the infants were nevertheless sensitive to the S/D relation. Furthermore, this experiment provides stronger evidence than the experiments of Tyrrell and colleagues where in each case there were only a very small set of items, making it unclear that a genuinely abstract relation was being detected.

Finally, it should be emphasized that there was no differences between the conditions in terms of either overall looking time or in decline and recovery of looking during the test phase. It might be expected that slides with two different objects are intrinsically more interesting than slides with two identical objects and that therefore infants in the *Different* condition would look longer, especially at the start of the experiment. In fact, no such differences were found. This is strong evidence that the 8 month old infants were responding on an abstract level to the S/D relation.

In contrast, the 4 month old infants did not show a significant decline in looking over familiarization and consequentially did not show the recovery or discrimination at test. They did not look longer at pairs of items exhibiting a novel S/D relationship relative to that which they had been familiarized with. One possibility is that the infants were potentially capable of making

the discrimination but because they had not habituated they were not liable to show a novelty preference at test. In practical terms it may be very difficult to test this hypothesis. The infants in this experiment, saw an average of over 18 habit trials each of up to 20 seconds and did not habituate. Either increasing the number of familiarization trials or increasing the maximum length of individual trials might give the infants more time to process the stimuli but just as likely is that the experiment would then be too long and lose the participants completely.

3 Experiment 2 - S/D Learning and Generalization using Anticipatory Eye Movement

This experiment adapts the anticipatory eye movement (AEM) paradigm developed by McMurray and Aslin (2004), building on the method of their Experiment 3. A pair of geometric shapes moved together behind an inverted T-shaped occluder and reappeared on either one side or the other. If the shapes were the same as each other they moved in one direction, if they were different they moved to the other direction. The experiment sought to determine if human infants would learn this dependency and correctly anticipate the reappearance of the shapes. A final phase tested if they could generalize this learning to the novel shapes.

The current experiment makes several methodological changes to the McMurray and Aslin paradigm. First, we reduced the speed with which the stimuli moved across the screen. In the original experiment stimuli were fully occluded

for only 750ms and infants appeared to be equally likely to expect the stimuli to appear along a linear trajectory directly above the occluder as they were to have looked either left or right, consistent with the actual paths the stimuli took. It is possible infants have a pre-potent expectation of linear motion (see von Hofsten et al., 2000) and so increasing the length of the occlusion allows time for the infants to overcome any such bias and search in the correct location. Therefore, in the current experiment the stimuli move more slowly. In particular, they remain occluded for 3000ms, with the whole animated sequence lasting approximately 6000ms. An additional benefit is that slower moving stimuli that remain on the screen for longer allow more opportunity for the infants to encode the items that are presented (Day and Burnham, 1981). A further problem of the original study was that infants missed some trials because they were not looking in the right direction. For the infant to learn the dependencies it is essential that they are attending to the screen. Therefore, in the present experiment, the software only started each trial when it determined that the infant was looking at the appropriate lower middle third of the screen. Finally, the number of learning trials is increased to give infants the fullest opportunity to learn the relation. The full procedure is explained below.

3.1 Method

3.1.1 Participants

The same infants who took part in Experiment 1 also took part in Experiment 2. Seventeen 8-months-olds were tested but only 10 infants (M = 252 days, range 243-259 days,) provided data suitable for group analysis. There were eyetracker timing and synchronization problems with data from 3 other infants and the remaining 4 provided insufficient data due to fussiness. A group of 9 four month old infants also provided data (M = 123 days, range 117-128 days, 6 female). A further 8 were tested but were excluded due to fussiness (3), synchronization problems in the eye-tracking data (3) or very sparse data due to poor calibration (2).

3.1.2 Apparatus

The experiment used a Tobii 1750 eye-tracking camera with integrated 17" LCD monitor. It was connected to a Dell PC running Windows XP and Exbuilder control software. A digital video camera placed directly above the monitor also recorded the infants reactions.

3.1.3 Stimuli

The stimuli are shown in Figure 3 and were seen by all the infants. An inverted T-shaped red occluder on a black background was present on the screen throughout each block. The occluder was approximately 23.3cm across and

11.6cm high at its maximum, subtending a visual angle of 21.9 ° by 11 ° when viewed at 60cm. For the learning phase the paired stimuli were a yellow circle and a light blue square. For the generalization phase, test shapes included a maroon cross, an orange heart, a pink diamond and a green triangle. Each was about 2.5cm across, making a visual angle of approximately 2.4 °.

3.1.4 Procedure

The infant was seated on the caregiver's lap in front of the Tobii 1750 at 60cm from the screen. A preliminary eye-tracking calibration was performed in which a small looming circle appeared at centre of the screen and moved sequentially to each corner. The experiment itself then immediately began. It was divided into four blocks, each of sixteen trials. The first three blocks were learning trials and featured only two of the stimulus shapes in all four possible pairings; two yellow circles or two blue squares, a circle and a square, or a square and a circle. Each pairing occurred four times in each block of sixteen learning trials with a randomized ordering for each infant and each block. In the generalization phase, eight novel pairings were introduced, four of which featured a maroon cross and orange heart in all possible pairings and four with a pink diamond and a green triangle. These were randomly mixed with eight trials that featured the circle and square from the learning phase. For half the infants all *Same* pairs would always go to the left and all Different pairs go to the right. This dependency was reversed for the other infants.

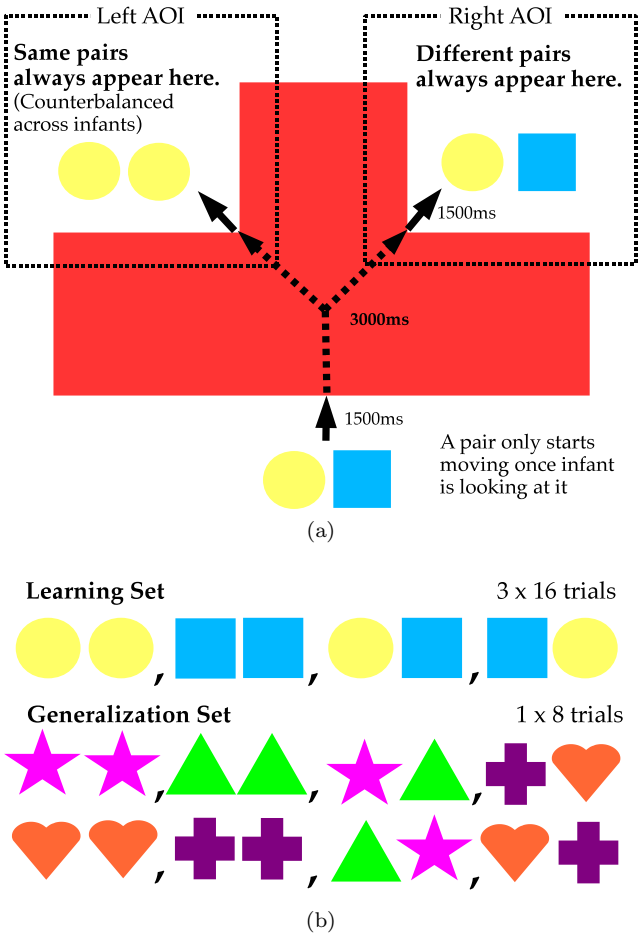


Fig. 3: (a) A schematic representation of the AEM task. On each trial two items would appear together at the bottom of the screen, move up and behind the inverted T-shaped occluder before reappearing on either the left or right depending on the S/D relation. The arrows and dotted squares were not visible but represent the path of the objects and the areas of interest used in the analysis respectively. (b) The pairs of shapes presented during the learning and generalization phases.

The procedure is illustrated in Figure 3. A bell sound effect marked the beginning of each trial when the two objects first appeared at the lower middle part of the screen. Once the eye-tracker detected the infant was attending to that part of the screen the objects would begin to move. Otherwise the

objects remained in place and loomed smaller and larger at a rate of about one full cycle per second until they captured the infants attention. The bell was then sounded again and objects began to move. In the first phase of each trial the objects moved together vertically upwards and behind the occluder, a transition which took approximately 1500ms. The objects then continued to move at the same rate following one leg of a Y-shaped trajectory behind the occluder and emerging on either the left or the right. The occlusion was accompanied by the sound of harp cascade. The objects were completely occluded for 3000ms. Finally the two objects moved to a resting position in the upper left or right when a final sound effect was played and the objects disappeared, either shrinking away to a point for all objects on one side or shrinking and spinning for all objects on the other side. This took a further 1500ms. The sound effects had an average intensity of around 73dB. To help the infants understand the task, on the very first four trials the occulder was semi-transparent so the infant could see the objects moving behind it. The occulder became gradually less transparent, so that from the fifth trial onwards it was completely opaque.

The experimenter attempted to get the infant to complete all 64 trials but this was rarely possible. Several steps were taken to increase the attentiveness of the infants. If necessary the experimenter could reorient the infant by causing randomly moving cartoon figures to appear on the screen between trials. Furthermore, in the third learning block the parents were instructed to talk to their children, making a 'peek-a-boo' style game out of the occlusion sequence.

They were encouraged to use phrases like “Where is it?”, “There it is!” that would maintain the infant’s engagement but without giving the infant any information about the trial outcomes. Caregivers were silent at all other times. Finally, if the experimenter felt the infant was still becoming fussy or otherwise disinterested during the final block of the learning phase, the experiment was moved directly into the generalization phase.

3.2 Results

In general the infants were engaged with the task. The 8-month-olds sat through an average of 43.2 out of 48 learning trials and 5.2 out of 8 generalization trials, the 4-month-olds saw an average of 37.7 learning trials and 4.9 test trials. Furthermore, there was no relation between the order in which experiments 1 and 2 were carried out and likelihood of a baby not completing the experiment [chi-squared: $\chi^2_1 = 1.83$, p = 0.18].

To compensate for the variation in the amount and quality of data collected only relative measures of performance were considered. Two equally sized areas of interest (AOI) were defined in the top left and top right and the analysis compared fixations to the correct AOI relative to those to the incorrect AOI. The two relative measures investigated were; (1) the proportion of fixation time during occlusion that was spent looking to the correct AOI and (2) whether the first fixation after occlusion was to the correct AOI. For the proportion score on each trial the cumulative fixation times to the correct AOI were divided by the total fixation to both AOI’s during the occlusion phase. For the first look

correct measure, each trial is scored as a success if during occlusion there was a fixation in the correct AOI *and* there was no prior fixation in the incorrect area, conversely if the first recorded look was to the incorrect AOI that trial is scored as a failure and otherwise the trial is marked as possessing insufficient data. All measures were calculated directly from the raw fixation data provided by the Tobii 1750.

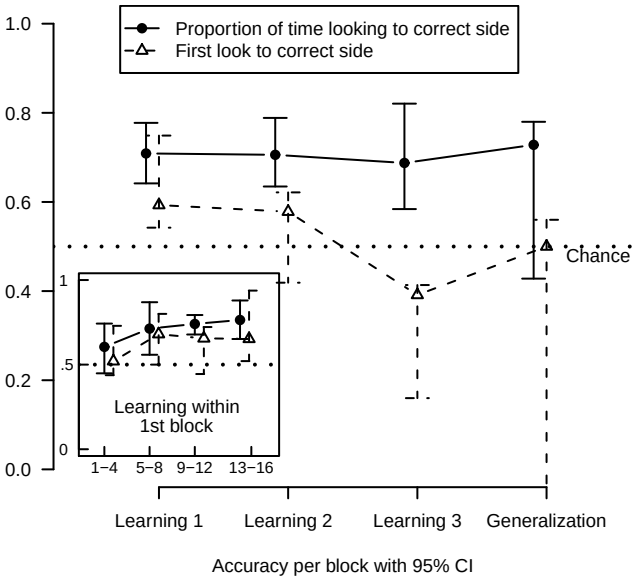
The sparseness of the data meant that not all participants had equal amounts of data per condition. Therefore, we opted to use weighted averages:

$$\mu^* = \frac{\sum_i w_i \bar{x}_i}{\sum_i w_i}$$

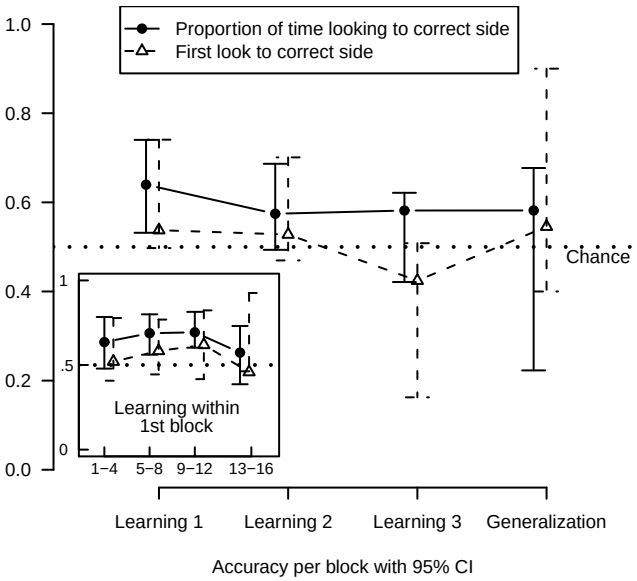
where, for a given participant in a given block, $\bar{x}_i$ is their average score and w_i is the number of data points collected for that participant. In this way, infants with more data-points contribute proportionally more to the estimate of the group mean but their individual (co)variance is ignored so that each participant is treated as an independent sample. However, using this measure means that standard parametric statistics should not be used. Instead, we used the bootstrap method of Efron (Efron and Tibshirani, 1997; Davison and Hinkley, 1997) to work out the 95% confidence intervals for each condition.[1]

The results organized by block are summarised in Tables 1 and 2 and shown in Figures 4(a) and 4(b). For the 8-month-olds, their relative looking to the correct area was significantly better than the chance level of 0.5 on

[1] Parametric analyses yielded similar results. Only results from bootstrapping are reported here.



(a) 8-month-olds



(b) 4-month-olds

Fig. 4: Accuracy scores for all participants in Experiment 2 averaged by block. The solid line represents the relative proportion of looking to the correct area of interest (AOI) during the occluded period. The dashed line represents the proportion of occlusions on which there was a fixation to the correct AOI *before* there was one to the incorrect AOI. The insets show learning within the first block. The error bars are 95% confidence intervals from the bootstrap analysis.

	Relative proportion looking to correct AOI		First look to correct AOI	
	Weighted Mean	95% Confidence Interval	Weighted Mean	95% Confidence Interval
Learning 1	0.71	0.64 – 0.78	0.59	0.55 – 0.75
Learning 2	0.71	0.64 – 0.80	0.58	0.42 – 0.63
Learning 3	0.69	0.59 – 0.82	0.39	0.16 – 0.41
Generalization	0.73	0.44 – 0.78	0.50	0.00 – 0.56
		Performance within the first learning block		
Trials 1–4	0.60	0.46 – 0.75	0.52	0.43 – 0.73
Trials 5–8	0.71	0.56 – 0.87	0.68	0.50 – 0.80
Trials 9–12	0.74	0.68 – 0.79	0.66	0.44 – 0.72
Trials 13–16	0.76	0.65 – 0.88	0.65	0.52 – 0.92

Table 1: Experiment 2 - Accuracy scores for 8 month old infants

	Relative proportion looking to correct AOI		First look to correct AOI	
	Weighted Mean	95% Confidence Interval	Weighted Mean	95% Confidence Interval
Learning 1	0.64	0.53 – 0.74	0.54	0.49 – 0.73
Learning 2	0.57	0.50 – 0.69	0.53	0.47 – 0.71
Learning 3	0.58	0.42 – 0.62	0.42	0.15 – 0.51
Generalization	0.58	0.21 v 0.68	0.55	0.40 – 0.90
		Performance within the first learning block		
Trials 1–4	0.64	0.48 – 0.79	0.52	0.43 – 0.78
Trials 5–8	0.69	0.55 – 0.80	0.58	0.45 – 0.77
Trials 9–12	0.69	0.60 – 0.82	0.62	0.40 – 0.82
Trials 13–16	0.57	0.41 – 0.73	0.46	0.46 – 0.91

Table 2: Experiment 2 - Accuracy scores for 4 month old infants

all three learning blocks. They maintained a similar level of accuracy in the generalization phase, although this was not significant. This may be a power issue as a consequence of there being less generalization data; an average of 5.3 points per participant compared to 16, 15.5 and 11.7 points in three learning blocks. Eight month old infants appear to have learned from the within the first block to associate the training stimuli with the sides they would reappear. Likewise, in the first block the infants would on average look first to the correct side significantly more often than chance. But this decreased over time and in

third learning block infants were looking more to the other side. This finding may reflect infants learning the time structure of the events. Objects always reappeared 3000ms after they were occluded so there is no need to look towards the AOI before that point and consequentially infants may learn to cast their first looks elsewhere. Although the 4-month-olds showed a qualitatively similar pattern of response. The 4-month were only significantly above chance in the relative looking for the first block. But, as before this may be an issue of power, as there was much less data for the age group. with 12.6, 10.6, 5.2 and 2.9 data points per participant for each of the three learning blocks and the test block respectively.

It is notable that at both ages learning happened very quickly. Even if we discount the first 4 trials where the occluder is partially transparent and infants could potentially see the path of the pairs. In order to establish that infants really are learning the task, we performed another analysis looking at the time course of learning in the first block. We divided it into 4 sub-blocks and calculated the weighted average for each group of 4 trials. We also ran another set of bootstrap analyses. The results are shown in the lower halves of Table 1 and 2 and in the inset boxes in Figures 4(a) and 4(b). These analyses clearly indicate that the infants at both ages are learning something about the task very early on.

One further analysis investigated if infants learn differently between the Same and the Different trial types. The trials were classified accordingly and another set of weighted averages and bootstrapped confidence intervals were

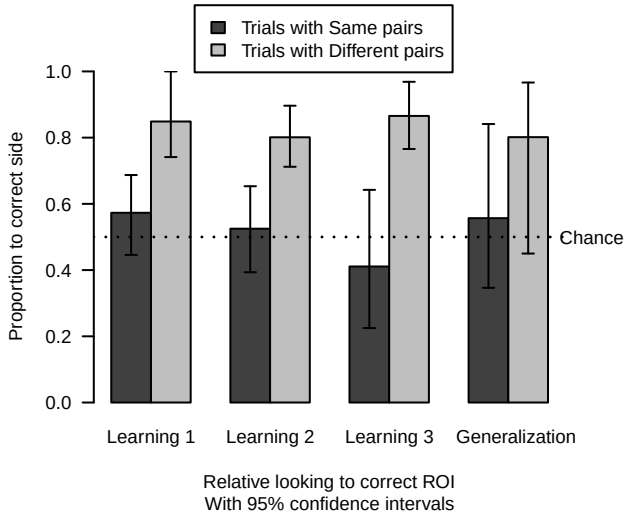
	Same pair trials		Different pair trials	
	Weighted Mean	95% Confidence Interval	Weighted Mean	95% Confidence Interval
Learning 1	0.57	0.45 – 0.68	0.85	0.74 – 1.00
Learning 2	0.52	0.39 – 0.66	0.80	0.71 – 0.90
Learning 3	0.41	0.20 – 0.65	0.87	0.76 – 0.97
Generalization	0.56	0.36 – 0.82	0.80	0.45 – 0.97

Table 3: Experiment 2 - Same versus different accuracy for 8 month olds

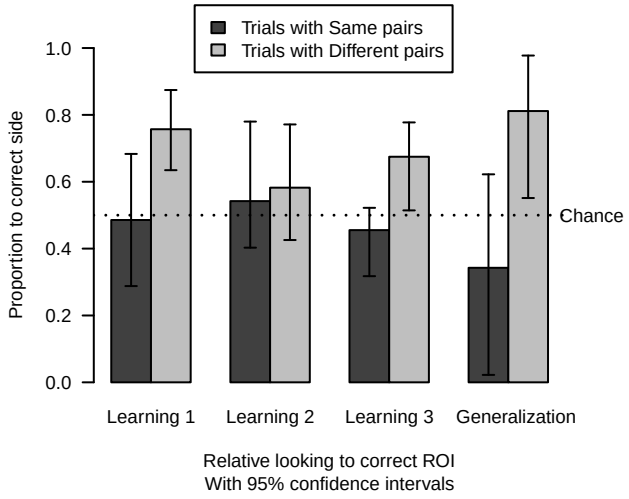
	Same pair trials		Different pair trials	
	Weighted Mean	95% Confidence Interval	Weighted Mean	95% Confidence Interval
Learning 1	0.49	0.28 – 0.69	0.76	0.64 – 0.87
Learning 2	0.54	0.40 – 0.80	0.58	0.42 – 0.78
Learning 3	0.46	0.33 – 0.52	0.67	0.52 – 0.77
Generalization	0.34	0.02 – 0.62	0.81	0.55 – 0.99

Table 4: Experiment 2 - Same versus different accuracy for 4 month olds

calculated. For this analysis we only looked at the more reliable proportion score. The results are presented in tables 3 and 4 and plotted in Figures 5(a) and 5(b).. The infants appear to learn to predict where the pairs of different items will reappear but are at chance when the items are the same as each other. To attempt to confirm this we ran a conventional analysis of variance on the combined learning data from the 4- and 8-month-olds using standard arithmetic means as the dependent variable, with age as a between subject variable and pair relation (same versus different) as a within subject variable. This revealed a main effect of age $F_{1,17}$ =5,76, p = 0.03 and a main effect of the S/D relation $F_{1,17}$ = 12.28, p = 0.003 but no other effects or interactions. This indicates that, although the effect was stronger with the 8-month-olds, both groups showed the same pattern of learning in only the different case.



(a) 8-month-olds



(b) 4-month-olds

Fig. 5: Graph showing the accuracy for same or different trials in the learning and generalization phases for Experiment 1b. Error bars show the 95% confidence intervals.

3.2.1 Comparing Experiments 1 and 2

Finally, we were interested in if there was any consistency in responses across the two experiments. As a measure of performance on the looking time study, we took the ratio of the two test conditions. A higher value indicates a stronger S/D response, namely a greater relative looking to the novel relation. For the AEM task we considered two measures, the average accuracy rate across the whole learning phase and accuracy in the generalization block as these may index different aspects of the ability. Since we were looking for evidence of common causation between the tasks we include both age groups in one analysis. This gave N=18 (10 eight-month-olds). The analysis showed that better performance in the looking time task did not correlate with performance in the AEM task, measured against learning (Spearman's rank correlation $\rho = 0.02$, p = 0.95) or generalization (Spearman's rank correlation $\rho = 0.19$, p = 0.50).

3.3 Discussion

The results of this experiment suggest that 8 month old human infants are able to use at least some aspect of S/D distinction on an instrumental learning task with simple geometric stimuli. It appears that these infants rapidly learn which side the different pairs will reappear and look longer to this location. The infants can generalize the rule they have learned, performing as well with 8 stimulus pairs with novel colors and shapes in the test phase. The infants' first

look after occlusion is initially predictive of where the objects will reappear but as they learn the timing of the reappearance they stop looking directly to the correct location. The fact that each trial had the same sequence of sound effects could also provide scaffolding for the infants to anticipate the point at which the objects reappeared. However, the fact that infants only appear to learn to predict the *Different* trials was unexpected and suggests that their performance may not be directly attributable to ability to perceive the abstract S/D relation, but may be due to a more low level interpretation.

The fact that infants only learned the dependency for *Different* trails but were at chance with *Same* pairings indicates that infants are responding to something other than a truly abstract S/D concept which in theory would lead to equal accuracy on both types of trials. From an adult perspective it is obvious that *Same* and *Different* are 'opposites' and so might be expected to be associated with different outcomes but this is a highly conceptual interpretation. The infants mixed performance makes it apparent that the task is not quite as simple as adults see it. In fact there are two separate contingencies to be learned. Infants must learn that non-matching pairs will always reappear on one side and, independently, that matching pairs always reappear on the other. In fact, as both Young and Wasserman (2002a) and Smith et al. (2008) make clear, there is a great deal of asymmetry between the two cases. Like the pigeons in Blaisdell and Cook (2005), infants could be responding to lower level aspects of the stimuli, symmetry, contrast, etc. that are present in the *Different* stimuli but not in the *Same* pairs. However, unlike Young and

Wasserman (2002a) the infants appear to have learned better with the more complex exemplars. This might be precisely because the *Different* stimuli are more complex and are therefore more engaging for the infants prompting them to learn. Alternatively, as suggested by Smith et al. (2008), the asymmetry may arise from the fact that a *Same* category is narrow and precisely delineated while *Different* is more broad. Hence *Different* is potentially easier for the infants to learn.

Interestingly, learning was very rapid. Infants were already above chance in the first block of 16 trials. Learning appeared faster than in Experiment 3 of McMurray and Aslin (2004), which used a very similar AEM design. In their learning phase infants received just 16 learning trials and averaged only 53% accuracy. Three factors might account for this improved learning (1) infants in our experiment were older (8 months vs 6 months), (2) each trial only started when the infant was looking towards the lower middle of the screen ensuring that all infants saw (at least part of) all the training trials, and (3) the stimuli moved more slowly across the screen giving infants more time per trial on being and included being fully occluded for 3000ms (vs 750ms) giving them more time to direct their attention to the correct location. This rapid learning can be found in other studies, with an alternative anticipatory eye movement design. For example, Sobel and Kirkham (2006, 2007) found that 8 and 5 month old infants were able to learn a spatial dependency after just 4 exposures to the critical training event.

The 4 month old infants in Experiment 2 did not show learning to the same extent as the 8 month old infants had done. However, they do seem to show a qualitatively similar pattern of responding. They learn the contingency for *Different* pairs but not for *Same* pairs and they generalize this learning to novel shapes. The fact that such young infants respond in this similar fashion supports our finding with 8-month-olds. When considered in light of their failure on the S/D looking time task, these results suggest that success in this task may be mediated by something other than abstract same-different concept learning. We return to this point in the general discussion below.

4 General Discussion

This study investigated the emergence of abstract concept learning in human infants. The two item S/D relation was chosen as an example of a very simple abstract concept that had previously been extensively investigated in the animal cognition literature but to a much lesser extent with human infants. Infants were tested in a passive looking time task and an active anticipatory eye movement (AEM) task with visual rich and simple stimuli respectively. Two groups of infants, aged 4 and 8 months, were tested to provide a developmental perspective. In the looking time task, it was found that 8-month-olds but not 4-month-olds were sensitive to the S/D relationship. In the AEM task, it was found that both 4- and 8-month-olds responded in a similar fashion; learning to anticipate the re-emergence of occluded Different paired shaped but performing at chance when the pair of shapes were the same as each other.

Additionally, both age groups transferred their learning in the Different case to trials with novel colored shapes. This pattern of responding is not consistent with with infants having a full awareness of an abstract concept of S/D but is strongly suggestive that the S/D concept develops from sensitivity to more perceptual regularities. Furthermore, no correlations were found between performance on the two tasks suggesting multiple processes are involved.

4.1 Looking times with photographic stimuli

In Experiment 1 infants were familarized with pairs of photographic images of either matching or non-matching inanimate objects drawn from a set of randomly chosen unrelated objects. They were then shown two further trials in which novel objects with a novel relation were presented. The results were very clear cut, with 8-month-olds succeeding and 4-month-olds failing at the task. In both conditions, 8-month-olds dishabituated to pictures showing the opposite relation to that which they had been familiarized with. There was no effect of scene complexity, meaning that infants did not look longer at displays with 2 dissimilar items relative to 2 identical items either during familiarization or at test.

These results provides clear evidence that 8 month old infants are sensitive to the abstract relation of S/D. This finding gives stronger support for infants' S/D discrimination ability than previous research. It is more direct proof than the indirect conclusions concerning S/D awareness than may be drawn from 7.5 month olds success on artificial grammar tasks (Marcus et al., 1999; Saffran

et al., 2007). It is a more robust finding than those of Tyrrell et al. (1991) and Tyrrell et al. (1993) where only a small set of objects were used and paired preference paradigms meant that Same and Different pairs were simultaneously presented. In the present study infants saw a wide range of unrelated objects whose only common feature was the S/D relationship between them and they only saw a single pair of items at a time.

The design of this study also provides continuity with the findings of Wright and Katz (2006) concerning S/D learning in monkeys and pigeons. These results show infants are also able to make a S/D discrimination using photographic stimuli, even more rapidly and with smaller example sets. This continuity with the animal findings and the strength of the effect in 8 month old infants suggests the mechanism is more general than previously believed. Note that the animal participants were rewarded for learning while human infants' performance is a measure of spontaneous discrimination. Moreover, the failure of 4-month-olds suggests that this S/D discrimination is not an innate response but is due to concept learning. Wright and Katz (2006) found that increasing the number of exemplars improved performance of the pigeons and monkeys so a similar manipulation may lead to better S/D discrimination in the younger infants. However, the current design is probably at the limit for of attention for the number of familiarization trials for such a young age group. (Infants saw up to 19 familiarization trials each of up to 20 seconds, plus 4 test trials). An alternative approach might be to have many more shorter trials and repetitions. Mareschal et al. (2005) found that a regime of rapid, repeated

presentation (each trial 2 seconds long seen multiple times) led 4-month-olds to succeed in forming categories based on perceptual correlations that would otherwise not be found before 7 months old. A similar design may reveal a sensitivity of 4-month-olds to the S/D relation.

Seven month old infants can learn the AAB/ABA discrimination with synthesized syllables (Marcus et al., 1999) and with pictures of dogs and cats (Saffran et al., 2007) but do not learn directly with musical tones, timbres or animal sounds (Marcus et al, 2007). Saffran et al. (2007) suggest that stimuli must belong to a familiar category if infants are to notice the higher order relations between them. But, in this experiment, the stimuli did not form a single meaningful category and many of the individual stimuli were very likely to be completely unfamiliar to the participants. Nevertheless, 8-month-olds were sensitive to the higher order S/D relation between the items.

4.2 Anticipatory eye movement with colored geometric shapes

In Experiments 2 an AEM paradigm demonstrated that infants can rapidly learn to anticipate where a pair of geometrical shapes would reappear but only when the pair of objects were different from each other. Infants spent longer looking to the correct area of the screen while the objects were occluded. The infants were also able to generalize their learning to novel shapes that also shared the different relation. But their performance was always at chance when the two objects were the same as each other. This was a surprising finding and one that is difficult to explain fully without further investigation.

Nevertheless, the effect appears to be real since a very similar pattern of responding was found in both 4- and 8-month-olds. Moreover, the asymmetry is strongly suggestive of a non-conceptual interpretation of infants abilities.

Unlike in Experiment 1, where infants must learn just one relation and simply detect its violation, in Experiment 2 the task is more complex; infants must simultaneously develop categories of *Sameness* and *Difference* and associate them with two possible outcomes. It appears that the infants, at both ages, rapidly learn which side the different pairs will reappear and look longer to this location. They can generalize this rule to novel pairs of shapes. But they do not learn the *Same* relation. This may imply that *Difference* is a simpler concept for infant than *Sameness*. One possible reason may be that to establish *Sameness* one must check that *all* properties of two objects are equivalent, whereas *any* single discrepancy establishes *Difference*, as suggested by Smith et al. (2008). Hence, under the challenging conditions of the AEM task, infants only manage to learn the simpler contingency. This difference between tasks and between *Sameness* and *Difference* within Experiment 2 shows that infants do not have a single, unitary S/D concept directly equivalent to that of adults.

One simpler explanation that has to be ruled out first is the possibility that infants' performance could be accounted for because they were exhibiting a systematic side bias, as was found in some infants in the experiments of McMurray and Aslin (2004). This can be discounted on two grounds. Firstly the side to which the Same and Different pairings appeared was counterbalanced

837 across infants, for half the infants Different pairs would go to the left and Same
838 pairs would go to the right, whilst for the remaining infants this dependency
839 was reversed. Therefore, nothing in the geometry of the experimental set up
840 (e.g., the layout of the testing room) could systematically influence infant at-
841 tention to one side or another. Secondly, a side bias would lead to a below
842 chance performance on the unbiased side. Such a bias was never seen in the
843 data with performance on the Same trials being very close to the chance level
844 of 50% in all cases.

845 Instead, it may be due to the infants responding to lower level aspects of
846 the stimuli such as symmetry, contrast, complexity, etc.. When Blaisdell and
847 Cook (2005) demonstrated two item S/D learning with geometric shapes in
848 pigeons they were careful to chose a set of colors and shapes for their stimuli
849 that would be highly dissimilar so that it would not be possible to account
850 for transfer in terms of the similarity between stimuli. They also point to
851 converging evidence from other studies that supports the case for an abstract
852 S/D in the pigeons by ruling out any other common perceptual property across
853 experiments (e.g. Cook et al., 1997; Cook, 2002). However, in Experiments 2
854 an S/D interpretation is *not* appropriate since there is no learning of the Same
855 cases and so alternative mechanisms must be considered.

856 In Young and Wasserman (2002a), pigeons learned S/D discriminations
857 better with larger arrays of items that increased their awareness of the redun-
858 dancy / dissimilarity contrast. Although those findings do not bear directly on
859 the performance of the human infants in this experiment, one related expla-

nation might be that the Different stimuli are more complex with contrasting colors and they were more engaging leading the infants to attend and learn with these stimuli. Although this interpretation seems implausible from an adult perspective, only further experiment could rule out such an explanation of the infant behavior. It is even more surprising that neither group of infants learned the Same dependency when the learning phase used only two objects (the yellow circle and the blue square). With just four trial types in the learning phase, it might be expected that infants could learn a direct association between each pairing and its eventual reappearance. However, this did not happen with either age group and the fact that the identical pattern was found in the generalization phase suggest that performance must be accounted for in terms of some kind of abstract feature of the stimuli. Future studies could systematically vary the common and distinct properties of the stimuli to investigate this. For example, using arrays of multiple objects or a wider selection of objects in the learning trials or using photographic stimuli to investigate the issue of the infants' engagement with the stimuli. Without further studies, it is not possible to give a clear explanation for the surprising findings of Experiments 2. Nevertheless, the identical pattern of responding between ages and between learning and generalization makes it likely that there is some abstract process underlying the infants behavior in this task.

880 4.3 Infants' understanding of Same/Different as an abstract concept?

881 While the results of these experiments show that 8-month-olds (but not 4-
882 month-olds) respond to the relation of Same and Different with photographic
883 stimuli and that both 4- and 8-month-olds learn and generalize some property
884 of Different geometric shapes, questions remains as to the extent to which these
885 abilities can be said to be *abstract* and the extent to which they can be said to
886 be *conceptual.* The two are not necessarily synonymous and the question is of
887 interest to developmentalists. Jean Mandler (Mandler, 2004) argues that there
888 is a fundamental difference between the perceptual and the conceptual. The
889 former referring to implicit perceptual sensorimotor knowledge while "concept
890 refers to declarative knowledge about object kinds and events that is poten-
891 tially accessible to conscious thought." (p.4, Mandler, 2004). In particular she
892 claims that concepts are not just complex associations of percepts but are a
893 different mode of representation created from perceptual inputs and a core set
894 of innate *image-schemas* (see chapter 4, Mandler, 2004). By contrast, Quinn
895 and Eimas (2000) take a strongly empiricist perspective arguing for continuity
896 between percepts and concepts. Our results help clarify this debate.

897 From Mandler's perspective there can be two interpretations of S/D dis-
898 crimination abilities found in these experiments. Either these abilities are due
899 to the infants detecting perceptual regularities in the stimuli they encountered
900 or could be a consequence of their awareness of a concepts of Sameness and
901 Difference. Since, in Experiment 1, the stimuli are all unrelated, it is hard to
902 point to the perceptual features that the stimuli have in common and which

would contribute to a concept of sameness. Sameness appears to be a concept but not a category. Here, the commonality between instances is of a different kind to that found in object-based categories such as trees or dogs. Infants as young as 3 months can form such categories (Eimas and Quinn, 1994) as can pigeons (Herrnstein and Loveland, 1964). In such cases, although it may not be possible to determine exactly which features define the category, it is clear that they are perceptually grounded. Members of these categories share many similarities in terms of texture, color, shape and spatial arrangement of features. Whilst all the Same pairs do share a horizontal translational symmetry, this is not a shared perceptual feature per se, but operates more like an abstract rule applicable to those cases. Nevertheless, Sameness or object-identity without being a concrete or abstract object-category is fundamentally a perceptual property and may contrary to Mandler's view develop from perception alone. A parallel can be found in Quinn's (2003) account of development of spatial relation categories in infancy. Quinn reviews evidence that infants under 1 year do have non-object based concepts of spatial relations such as above, below, between but that these develop in sophistication and abstraction over the first year consistent with a perceptually grounded account, where innate core concepts are not necessary. The failure of 4-month-olds to respond to the S/D distinction with photographic stimuli indicates that this ability also develops. Although this by itself does not necessarily support either position.

The results of Experiment 2 also support a perceptual interpretation. Infants are not applying an abstract S/D rule to the stimuli but are abstracting

some perceptual feature to identify the Different stimuli and to learn to predict where these stimuli reappear and generalize this learning to new Different stimuli. This ability is largely unchanged across the two ages tested who show a very similar pattern of response. Infants of both ages seeming to possess the same ability to discriminate and generalize only with Different stimuli. A surprising finding made more so when we compare this to the successes and failures of the same group of infants on the looking time tasks. A result which is problematic for Mandler's dual process account; if eight month old infants possessed an abstract conceptual awareness of the S/D relation that allowed them to succeed on the looking time task, why is this concept not accessible in the AEM task? If the abilities are based on a perceptual analysis then this difference between the tasks can be explained in terms of the substantial perceptual differences between the stimuli used in the task. This interpretation leads to the testable predictions that using simpler stimuli in the looking time task might lead 4-month-olds to succeed on the S/D discrimination whilst using more complex stimuli in the AEM may induce the older infants to learn the dependencies for the Same stimuli.

Of course, there is a further problem with a conceptual interpretation to the looking time results with 8-month-olds in that monkeys and pigeons can also make S/D discriminations with photographs (Wright and Katz, 2006). Likewise, Blaisdell and Cook (2005) show that pigeons can discriminate S/D relations with colored geometric shapes. In neither case would any researchers be likely to attribute these successes to the animals possessing human-like

concepts but these results do prove problematic for Mandler's account, where perceptual meaning analysis is required to form and utilize abstract concepts. Mandler (2004) is dismissive of Young and Wasserman's (2002) results, claiming that these are purely abstract perceptual abilities but does not address the wider comparative literature where the accumulation of data across a range of paradigms make clear that non-human animals can go beyond the purely perceptual features of S/D stimuli. Penn et al. (2008) do look at a wider set of animal studies but make their strong claims that human and non-human S/D abilities are qualitatively different without considering how human infants perform on these tasks. Our results bridge this gap and support the position of Wasserman and Young (2010) that there are continuities between humans and animals. We find that infants show asymmetries between *Same* and *Different* and that, as in other species, their abilities depend upon the particulars of experimental set-up and perceptual features available in a given task. The S/D concepts of both human infants and animals appear to be firmly grounded in perception.

Acknowledgements We wish to thank Dick Aslin for useful discussion of the AEM paradigm, Edward Longhurst for support using the Exbuilder software and Fani Deligianni for assistance with the implementation. We thank several anonymous reviewers for the helpful comments on early draft of this paper and we particularly acknowledge the suggestion from Bob McMurray of using bootstrap methods.

This research was approved by the research ethics committee of the Department of Psychological Sciences, Birkbeck, University of London and complied with the current laws

of Great Britain, where it was carried out. The authors declare that they have no conflict of interest.

References

Blaisdell AP Cook RG (2005) Two-item same-different concept learning in pigeons. Learn Behav 33:67–77

Brainard DH (1997) The psychophysics toolbox. Spat Vis 10:433–436

Caron AJ Caron RF (1981) Processing of relational information as an index of infant risk. In: Friedman S Sigman M (eds), Preterm birth and psychological development., Academic Press, New York, pp. 219–240

Castro L Young ME Wasserman EA (2006) Effects of number of items and visual display variability on same-different discrimination behavior. Memory and cognition 34:1689–1703

Cook RG (2002) The structure of pigeon multiple-class same-different learning. J Exp Anal Behav 78:345–364

Cook RG Katz JS Cavoto BR (1997) Pigeon same-different concept learning with multiple stimulus classes. J Exp Psychol Anim B 23:417–433

Davison AC Hinkley DV (1997) Bootstrap methods and their application. Cambridge University Press, Cambridge

Day RH Burnham DK (1981) Infants' perception of shape and color in laterally moving patterns. Infant behav dev 4:341–357

Efron B Tibshirani RJ (1997) An introduction to the bootstrap. Chapman and Hall, Boca Raton

995 Eimas PD Quinn PC (1994) Studies on the formation of perceptually based
996 basic-level categories in young infants. Child Dev 65:903–917
997 Flemming TM Beran MJ Washburn DA (2007) Disconnect in concept learning
998 by rhesus monkeys (*Macaca mulatta*): Judgment of relations and relations-
999 between-relations. J Exp Psychol Anim B 33:55–63
1000 Giurfa M Zhang S Jenett A Menzel R Srinivasan MV (2001) The concepts of
1001 'sameness' and 'difference' in an insect. Nature 410:930–932
1002 Herrnstein RJ Loveland DH (1964) Complex visual concept in the pigeon.
1003 Science 146:549–551
1004 House BJ Brown AL Scott MS (1974) Children's discrimination learning based
1005 on identity or difference. In: H.W.Reece (ed), Advances in child development
1006 and behavior, vol. 9,. Academic Press, New York, pp. 1–45
1007 Johnson SP Fernandas KJ Frank MC Kirkham NZ Marcus G Rabagliati H
1008 Slemmer JA (2009) Abstract rule learning for visual sequences in 8-and
1009 11-month-olds. Infancy 14:2–18
1010 Mandler JM (2004) The Foundations of Mind - Origins of Conceptual
1011 Thought. Oxford University Press, Oxford
1012 Marcus G Fernandes K Johnson S (2007) Infant rule learning facilitated by
1013 speech. Psychol Sci 18:387–391
1014 Marcus GF Vijayan S Bandi Rao S Vishton PM (1999) Rule learning by
1015 seven-month-old infants. Science 283:77–80
1016 Mareschal D Powell D Westermann G Volein A (2005) Evidence of rapid
1017 correlation-based perceptual category learning by 4-month-olds. Infant

1018 Child Dev 14:445–457

1019 Mareschal D Quinn PC (2001) Categorization in infancy. Trends Cogn Sci
1020 5:443–450

1021 Mareschal D Quinn PC Lea SEG (eds) (2010) The making of human concepts.
1022 Oxford University Press, Oxford

1023 McMurray B Aslin RN (2004) Anticipatory eye movements reveal infants'
1024 auditory and visual categories. Infancy 6:203–229

1025 Mercado E Killebrew DA Pack AA Mácha IVB Herman LM (2000) Gener-
1026 alization of 'same-different' classification abilities in bottlenosed dolphins.
1027 Behav Processes 50:79–94

1028 Oden DL Thompson RKR Premack D (1990) Infant chimpanzees sponta-
1029 neously perceive both concrete and abstract same/different relations. Child
1030 Dev 61:621–631

1031 Pearce JM (1988) Stimulus generalization and the acquisition of categories
1032 by pigeons. In: Weiskrantz L (ed), Thought without language, Clarendon,
1033 Oxford, pp. 132–152

1034 Pearce JM (1991) The acquisition of concrete and abstract categories in pi-
1035 geons. In: Dachowski L Flaherty C (eds), Current topics in animal learning:
1036 Brain, emotion, and cognition, Erlbaum, Hillsdale, NJ, pp. 141–164

1037 Pelli DG (1997) The videotoolbox software for visual psychophysics: trans-
1038 forming numbers into movies. Spat Vis 10:437–442

1039 Penn DC Holyoak KJ Povinelli DJ (2008) Darwin's mistake: Explaining
1040 the discontinuity between human and nonhuman minds. Behav Brain Sci

31:109–130

Premack D (1983) Animal cognition. Annu Rev Psychol 34:351–362

Quinn PC (2003) Concepts are not just for objects. In: Rakison D Oakes L (eds), Early Category and Concept Development, Oxford University Press, Oxford, pp. 50–76

Quinn PC Eimas PD (2000) The emergence of category representations during infancy: are separate perceptual and conceptual processes required? J Cogn Dev 1:55–61

Saffran JR Pollak SD Seibel RL Shkolnik A (2007) Dog is a dog is a dog: infant rule learning is not specific to language. Cognition 105:669–80

Smith JD Redford JS Haas SM Coutinho MVC Couchman JJ (2008) The comparative psychology of same-different judgments by humans (*Homo sapiens*) and monkeys (*Macaca mulatta*). J Exp Psychol Anim B 34:361–374

Sobel DM Kirkham NZ (2006) Blickets and babies: the development of causal reasoning in toddlers and infants. Dev Psychol 42:1103–1115

Sobel DM Kirkham NZ (2007) Bayes nets and babies: infants developing statistical reasoning abilities and their representation of causal knowledge. Dev Sci 10:298–306.

Tyrrell DJ Stauffer LB Snowman LG (1991) Perception of abstract identity/difference relationships by infants. Infant behav dev 14:125–129.

Tyrrell DJ Zingaro MC Minard KL (1993) Learning and transfer of identity - difference relationships by infants. Infant behav dev 16:43–52.

1063 von Hofsten C Feng Q Spelke ES (2000) Object representation and predictive
1064 action in infancy. Dev Sci 3:193–205
1065 Wasserman E Young M (2010) Same–Different Discrimination: The Keel and
1066 Backbone of Thought and Reasoning. J Exp Psychol Anim Behav Process
1067 36:3–22
1068 Wright AA (1997) Concept learning and learning strategies. Psychol Sci
1069 8:119–123
1070 Wright AA Katz JS (2006) Mechanisms of same/different concept learning in
1071 primates and avians. Behav Processes 72:234–254
1072 Young M Wasserman E (2001) Entropy and variability discrimination. J Exp
1073 Psychol Learn 27:278–293
1074 Young ME Wasserman EA (2002a). The pigeon's discrimination of visual
1075 entropy: a logarithmic function. Anim learn behav 30:306–314
1076 Young ME Wasserman EA (2002b). Detecting variety: What's so special about
1077 sameness? J Exp Psychol Gen 131:131–143
1078 Zentall TR Edwards CA Moore BS Hogan DE (1981) Identity: the basis for
1079 both matching and oddity learning in pigeons. J Exp Psychol 7:70–86
1080 Zentall TR Hogan DE (1974) Abstract concept learning in the pigeon. J Exp
1081 Psychol 102:393–398
1082 Zentall TR Hogan DE (1976) Pigeons can learn identity or difference, or both.
1083 Science 191:408–409